What Works When?
White Paper on Web Site Testing Methodologies

© Web Mystery Shoppers Inc. 2005

- "You can increase sales on your site as much as 225% by providing sufficient product information to your customers at the right time.… By understanding your customer expectations and needs, and designing your product lists accordingly, you can significantly increase your sales." *(UI Engineering, 2001)*

- "One study estimated that improving the customer experience increases the number of buyers by 40% and increases order size by 10%." *(Creative Good, 2000)*

- "The magnitude of usability improvements is usually large. This is not a matter of increasing use by a few percent. It is common for usability efforts to result in a hundred percent or more increase in traffic or sales." *(Nielsen, July 1999)*

- "Convoluted e-commerce sites can lose up to half of their potential sales if customers can't find merchandise, according to Forrester Research, Inc." *(Kalin, 1999)*

If you are considering revising your web site or putting one up in the first place, an important way to ensure you make a wise investment is to have independent user-testing done. You, your employees, your friends and family all know your business too well to give you an objective view of what it would feel like for a new prospect to visit your site. And your site designer knows computers too well to give you that realistic view. If your designer hits a snag on a web site, he or she undoubtedly knows the tricks of the trade to get around it. But your potential customers don't.

This paper discusses different approaches to getting feedback on your web site. In an ideal world, you would use each of the techniques discussed in this paper at some point during your web site's life cycle. But this isn't an ideal world, and odds are you don't have the time or the budget to do them all. So how do you decide?

The type of testing you do should depend on a combination of your stage in the web site life cycle and your goals for testing. Realistically, it will also be influenced by cost and time pressures.

This paper examines six web site testing methods:
1. Focus groups (in-person or online)
2. Lab-based usability testing
3. Web-based (pop-up) surveys
4. Metrics/Analytics software
5. Unsolicited customer feedback
6. Remote usability testing.

If you don't have the time to read all the details, flip to the summary chart on the last page.

Focus Groups

These typically involve 6 to 12 people, who may be chosen from your existing customer base, or recruited externally. They are brought into a meeting room where they engage in a group discussion, led by a moderator, about the web site. If the site does not yet exist, they may be asked what they would like to see in such a site. They may be shown page mock-ups, drawings, or led through an exercise called "card-sorting", which is intended to help the site designers figure out how users would interpret various topic labels, and how they would organize the information.

Often the client is behind one-way mirror, and the facility may be set up in such a way that the client can ask the moderator (who is wearing a headset) to ask specific follow-up questions. Costs are based on a combination of recruitment costs, participant expenses, and consulting-related fees that may be charged by the company running the focus group. Typical focus groups cost between $3,000 - $8,000 per group, which includes a fee of typically $50 to $100 per participant.

Advantages of Focus Groups

✓ Client gets to watch participant reactions, and may be able to have moderator ask follow-up questions.

✓ Participants can "play" with the content – for example, rearranging cards in an order that makes sense to them.

✓ Can be done early – before money has been spent on programming.

Drawbacks of Focus Groups

✎ User's guesses as to how they would use the site are very different from reality. Will miss most navigational problems.

✎ Risk of "groupthink"; what people say is influenced by the others in the room.

✎ Risk of poor moderators. Need a good moderator – must not bias the conversation, but ensure no one dominates.

✎ Participants limited by geography, time and willingness to travel to the testing location.

✎ Often either no transcript available, or significant cost or delay to get a transcript.

✎ Poorly constructed participant group negates the value (i.e. they must be like your target customers, and not so familiar with your site or company that they are unable to view the concepts with fresh eyes).

Focus Groups Useful When:

➢ In very early stages -- conceptual work.

➢ For emotional reactions to colours or treatments.

➢ To get rough idea of desired content.

On-line Focus Groups

A variant of traditional focus groups, on-line focus groups allow participants to log in from their own computers at a pre-arranged time, and the discussion uses a chat-room approach. In the more sophisticated set-ups, participants can view materials that the moderator displays on the screen, and possibly go back and forth between the chat room and the viewing screen, or see both at once.

Advantages of On-line Focus Groups

✓ Overcomes geographic limitations of traditional focus groups.

✓ Real-time transcript is available for analysis.

Drawbacks of On-line Focus Groups

↗ Cannot do physical activities such as card-sorting.

↗ Hard to follow the flow of conversation.

↗ No facial expressions or tone of voice.

On-line Focus Groups Useful When:

➢ Doing early-stage testing for a teen-oriented site. Teens tend to be the group that is most comfortable with on-line chat, and they may prefer the semi-anonymity of this format to a traditional focus group. For most other users, this is not an ideal format.

Lab-based Usability Testing

Users are brought into a testing lab, one at a time. A typical session lasts one to two hours with each user. The lab may be configured to resemble an office or a home. The participants work on the lab's computer (there may be a choice between a Mac-based OS and a Windows-based system), and are asked to vocalize their reactions ("think aloud").

The client can usually watch behind a one-way mirror. As with a focus group, there may be a "moderator" present in the lab with the user, who can prod for further insight, ask follow up questions, or, in extreme cases, give permission for the user to give up and try something else.

In some labs there is also a video-camera taping the participant's facial reactions, and some form of clickstream monitoring (tracking of where the person clicks or scrolls on a page).

Given the intensity of this type of testing, there are usually only a handful of people tested. The traditionally accepted minimum was five testers, but current research suggests that this is far too few for today's complex web sites. The problems identified by the first few users may not reflect the most important problems in need of repair.

A typical set of lab-based tests with 5 – 10 users will cost in the range of $20,000 - $100,000, depending on the degree of consulting that is included in the project and the market clout of the testing company.

Advantages of Lab-based Usability Testing

✓ Users are interacting with the real site, so you can learn about navigational problems.

✓ Can ask the users questions as they interact with the site, so surprising actions can be explained.

✓ Because you see the user's facial expressions and hesitations, you get a sense of the frustration level and uncertainties experienced by the user.

Drawbacks of Lab-based Usability Testing

↗ It is a stressful, unnatural environment. Research clearly demonstrates that stress affects performance.

↗ Using an unfamiliar computer, which will cause frustrations that the shopper would not encounter on his or her own computer.

↗ Most lab-based computers are more finely-tuned and using more current software than most users' real computers, so many errors users encounter in real life will be missed in this scenario.

↗ Users need to be willing and available to come to lab. This causes a skew in who is willing to participate.

↗ Geographic limitations. Most labs are in major cities. But are all your users?

↗ Small number of users tested. Are the problems the first 5 – 10 people encounter typical?

🖊 Often expensive, particularly if a transcript is prepared and analyzed after-the-fact.

Lab-based Usability Testing Useful When:

➢ Want to see facial reactions; hear vocalized thoughts.

➢ Once have at least a "draft" site available for on-screen use.

➢ Want to be able to intervene with real-time questions.

➢ Site users are a very homogeneous group, so it is possible to recruit a small, but representative sample.

Website (Pop-up) Surveys

There are many companies that offer surveys, which can either be mailed to site visitors after the fact, or that pop up on site while the visitor is using it. Mailed surveys are typically too long after the fact to yield any meaningful results for web site assessment. At most you will get vague, general feelings about the overall experience. ("It left me feeling frustrated." Or "It was OK".)

Pop-up surveys are programmed to either pop up for every visitor, or for every nth visitor, or for visitors to specific pages. Visitors are asked if they will participate in a survey, and sometimes offered a chance to win a prize as an incentive for participation. They survey asks a few questions about the site experience. Generally, a maximum of about five minutes worth of questions and answers is the longest you can expect people to stick with such a survey.

Costs vary from small, free surveys offered by companies such as Zoomerang, to sophisticated surveys that change the questions asked depending on previous answers and have a lot of built-in analytic tools. The latter type of survey can easily cost in the $10,000 - $20,000 range. Because a heavily trafficked site will usually be able to collect a large number of responses very quickly in this way, pricing is often quoted on a per-respondent basis.

Advantages of Website (Pop-up) Surveys

✓ Large numbers can be surveyed

✓ Fast turn-around

✓ Low cost/respondent

✓ Actual site visitors responding

✓ Questions can be changed often

✓ No need to pay respondents

Drawbacks of Website (Pop-up) Surveys

🖊 Survey fatigue: people are less willing than they used to be to spend time answering such surveys.

🖊 Bias in who replies. Respondents are more likely to be new Internet users, as they are less likely to have become accustomed to being asked their opinion on web sites.

🖊 Intrusive; it may annoy site users. (NOTE: One common mistake is to ask visitors for their opinion *before* they've had a chance to explore site!)

🖊 Increasing prevalence of pop-up stoppers is making it harder to get results, especially from experienced internet users.

🖊 Very limited number of questions feasible, so you cannot get in-depth responses.

🖊 No tracking of the path taken and the customer's reactions as they moved along that path.

🖊 Requires special coding be placed on your site. This places demands on IT.

Website (Pop-up) Surveys Useful When:

➢ Site is already live.

➢ Want general reactions to site.

➢ Have very large numbers of visitors.

➢ Want feedback quickly, but not in-depth.

➢ Your target market consists mainly of new Internet users.

Metrics/Analytics Software

Your web site hosting company should be able to provide you with some basic statistics about who is visiting your site: how many unique visitors, which pages they are visiting, where they came to the site from, etc. Metrics or analytics software consists of sophisticated versions of these server logs. Some can track clickstream details (what pages people clicked on and in what order) and analyse patterns and drop-off points. They can usually track which pages were viewed and the average length of time spent per page.

They are useful in conjunction with other methods, but you are left to infer customer motivations. They may tell you, for instance, that 63% of your site visitors abandon your site at the first page of your order form. You have no way of know whether they are leaving at that point because the page isn't displaying properly on the screens of Windows XP users, or if they couldn't find information about shipping costs and were unwilling to continue without knowing that, or if they pressed the wrong button, and then decided it wasn't worth the hassle to go back and re-enter the information, etc.

Costs for such packages range from the free basic statistics provided by many site hosting companies, to sophisticated programs in the tens or hundreds of thousands of dollars.

Advantages of Site Metrics/Analytics Programs

✓ Can provide detailed statistical information on your actual site visitors.

✓ Can help you identify problem pages.

Drawbacks of Site Metrics/Analytics Programs

✎ The volume of statistical data can be overwhelming. Many site owners find it impossible to interpret, and end up ignoring most of it.

✎ Does not explain *why* visitors are behaving as they are.

Site Metrics/Analytics Programs Useful When:

➢ You have the staff and budget to do sophisticated analysis of the data, and can combine it with qualitative information gained from other approaches.

➢ It is provided free along with your hosting agreement, if you use it in conjunction with other approaches.

Unsolicited Customer Feedback

Almost everybody gets some feedback on their site. From family and friends to site visitors who feel strongly enough that they'll take the time to send you an e-mail or fill in a comments form, people will give you their opinions. If you have a call centre, you can also learn a lot about peoples' reactions to your site from analyzing the help requests that relate to the web site.

Since you get it, use it. You may get some gems from this sort of feedback. The big mistake site owners often make, though, is assuming that what they hear from such channels is typical, and that they thus know all they need to

know about what people think of their site. Wrong! The people who take the time to comment are not typical of the general public. They are a vocal minority, who tend to perceive things in more extreme terms than most, and who have enough time on their hands that they consider it worth the time to comment.

Advantages of Unsolicited Customer Feedback

✓ Minimal cost.

✓ Real users.

✓ Chance to repair relationship with dissatisfied customers who bother to comment.

Drawbacks of Unsolicited Customer Feedback

✗ Biased sample.

✗ Miss many issues.

✗ Analysis can be costly in large organizations.

✗ You must respond quickly and well, or your lack of correct response will make matters worse.

✗ Often you will not learn exactly where the problem occurred. The comments are frequently too vague to be useful.

Unsolicited Customer Feedback Useful when:

➢ Making minor adjustments to the site as you go along.

➢ Ensuring a good linkage between customer service staff and site promises.

➢ Input for redesign.

Remote Usability Testing

A new approach is what's called "remote usability testing". Some people mistakenly call online focus groups remote usability testing, but, as discussed earlier, focus groups help you develop concepts but do not help learn about the actual usability of a web site.

One remote usability approach is to work one-on-one, as in a lab-based usability test, but with the testers working from their own computers and either talking by phone or through on-line chat with the interviewer while they work through the site. Typically this uses technology that allows the interviewer to see the same screen as the person doing the testing. Let's call this "One-On-One Remote Usability Testing". The advantages and drawbacks of this approach are similar to lab-based usability testing, except that you can use testers from any geographic region, but you cannot see their facial expressions. (Although sometimes these are videotaped and streamed back to the tester.)

Another approach is to have large numbers of people (typically 100 – 200) work independently on their own computers to try to complete tasks at the site, and report on their progress by answering survey forms as they go through the site. Since they are expected to devote much more time to the process than is expected in a pop-up survey, they are usually paid $10 - $20 for their effort. We'll call this "Large Sample Size Remote Usability Testing" (LSSRUT).

Advantages of Large Sample Size Remote Usability Testing (LSSRUT)

✓ No geographic or time-based limitations.

✓ Collect lots of detail as they go.

✓ Includes both qualitative and quantitative information.

✓ Realistic setting; low stress.

✓ Wide range of computer systems; uncover problems faced using buggy personal computers.

✓ No demands on IT department*

✓ Provides easy ability to benchmark – both against competitors and to compare present and future performance*.

    * (Depending on system used to collect the information)

Drawbacks of LSSRUT

✒ Imperfect click-stream analysis. The testers must report on where they clicked rather than having it tracked automatically.*

✒ Can't see facial expressions.

✒ Needs reasonably literate testers.

✒ Site needs to already be accessible on-line (even if only on a test server).

    * (Depending on system used to collect the information)

LSSRUT Useful when:

➢ Preparing for redesign

➢ Competitive analysis*

➢ Benchmarking*

➢ Monitor and reward customer service staff*

➢ Need either a wide range of testers (e.g. people from across North America) or a specific type of testers (e.g. rural Albertans aged 50+)

➢ A current site or a draft version of a new site is accessible on-line

    * (Depending on system used to collect the information)

Which Techniques are Best at What Stages?

Concept development: Use focus groups, plus feedback from clients and current customers.

Once online (test or live server): Use lab-based testing if you need to see facial expressions; otherwise consider alternatives such as LSSRUT.

Ongoing: Pop-ups if and when needed; web analytics/metrics, collect and analyze unsolicited feedback; use LSSRUT to ensure not slipping and for competitive analysis.

Redesign: LSSRUT _before_ redesigning, so you don't throw out the baby with the bathwater! Also use data from ongoing stage.

Comparison of Testing Techniques

	Lab-based Usability Test	Focus Group (live)	Focus Group (online)	Pop-up Surveys	Metrics / Analytics Software	Unsolicited Customer Feedback	LSSRUT*
Unbiased, representative sample					✓		✓
Realistic view of actual user paths	✓				✓		✓
Uncovers navigational problems	✓				✓		✓
Transcript available for details	with delay	with delay	✓				✓
No IT involvement to implement		✓					✓
No risk of group-think	✓			✓	✓	✓	✓
Uncover computer-related problems					✓		✓
Large sample size				✓	✓		✓
Test in customers' real environment				✓	✓		✓
Statistical data				✓	✓		✓
Qualitative data	✓	✓	✓			✓	✓
Can test & compare to competitors	✓						✓
Can be done before go live		✓				✓	✓
Can see reactions	✓	✓					
Can ask user follow-up questions	✓	✓	✓				

*Large Sample Size Remote Usability Testing. Not all LSSRUT approaches have all the benefits outlined here. It depends on how the usability testing company has structured its data collection process. For more information call Tema Frank at (780) 444-5645 or toll-free 1-866-211-7027, or send her an e-mail: tema@webmysteryshoppers.com